

## **Breadth of focus, modality and prominence perception in Neapolitan Italian\***

**Mariapaola D'Imperio**  
dimperio@ling.ohio-state.edu

**Abstract:** This study explores the notions of "nuclear stress", "accent placement" and "breadth of focus" in the Neapolitan variety of Italian. The predictions of standard generative theories about their interrelationships are tested through a perceptual study employing statements and questions with varying focus structure. The results show that broad focus statements are more ambiguous than late narrow focus ones as to the extraction of intended focus pattern. Broad focus questions are, in turn, less ambiguous than broad focus statements for the same purpose. The results suggest the importance of the role of accent type differences.

### **INTRODUCTION**

The goal of the present study was to test two claims about the relationships among prosodic prominence, intonational accent and scope of focus that are standardly assumed in generative linguistic theories. The language investigated is the variety of standard Italian spoken in Naples. Two experiments with forced choice response were designed to assess different aspects of this complex of phenomena. Experiment I was intended to elicit responses that indirectly measured location of prominent words, while Experiment II explored focus scope more directly through a question-answer matching paradigm. For both experiments, speech stimuli varying in focus placement (early, medial, late) and focus type (broad vs. narrow) were employed.

Theoretical claims about the relation between accent and stress on one side, and between focus and accent/stress on the other will be reviewed in order to formulate hypotheses and predictions. Therefore the next section will be concerned with these claims and with possible implications for the present study.

### **SENTENCE STRESS, NUCLEAR ACCENT AND FOCUS OF INFORMATION**

In describing English intonation, two main themes recur in the literature, namely the notion of "sentence stress" and the notion of "focus" and its scope. I shall first try to clarify the former notion as it has been interpreted in various theoretical frameworks.

Sentence stress was defined in the earliest versions of Generative Phonology as the derivational product of the "nuclear stress rule" as formulated by Chomsky and Halle (1968). In neutral utterances, the application of this rule renders the last stress of a phrase the most prominent, that is the "nuclear" stress of the utterance. The nuclear stress corresponds to the "tonic syllable" in works such as Halliday (1967) and to the "nuclear syllable" in Crystal (1969). In these theories, sentence stress was recognized to be inextricably tied to the intonation pattern; the nuclear stress was defined in terms of the

---

\* This work represents the partial results of a research that I started last summer, thanks to the Cognitive Science Department Summer Fellowship, which was inspired by my conversations with Prof. Mari Jones. Special thanks to Prof. Mary Beckman for extensive guidance in the whole process; thanks also to Mike Calcagno, Craige Roberts and all the members of the "Pragmatics" group for helpful suggestions regarding Experiment II.

location of the only prominent pitch event or "accent" of the intonational phrase. Pierrehumbert (1980) and subsequent works by her colleagues changed this view, in that in her theory the nuclear accent is not necessarily the unique prominent melodic event of the intonational phrase; rather, it is one of the potential accents of the phrase (though it is characterized by having a special status).

In the original Pierrehumbert's theory, which was cast in a derivational framework, stress assignment precedes accent assignment in that "pitch accents are lined up with the text on the basis of the prominence relations" (Pierrehumbert 1980, p.102). Therefore, a pitch accent must occur on a syllable that has already been designated by the grammar as being metrically the strongest, or most "stressed", in that phrase. Selkirk (1984) embodies another version of the traditional derivational approach, but in her framework accent assignment has to precede stress assignment. More recent approaches suggest that we do not need a serial process of stress-to-accent or accent-to-stress application, but we can think of various constraints acting in parallel to produce a well-formed output (Pierrehumbert 1993; Beckman, 1996).

The original theory was couched in terms of Liberman and Prince's (1977) account of stress as metrical (i.e. rhythmic) strength. It stated that the nuclear accent was associated to the D.T.E. (Designated Terminal Element) syllable, and that earlier accents could be associated only to relatively strong syllables, at grid levels no higher than the nuclear accented syllable. Beckman (1986) proposed a different account of metrical strength, which incorporated the results of phonetic experiments such as Fry (1958). It is this more intonationally "direct" account of stress that is incorporated in the ToBI conventions (Beckman & Ayers, 1994). In this account, stress is envisaged as a hierarchy of prominence levels, where each level is defined in terms of its own phonetic properties. It is useful to think of this idea in terms of the following grid model:

4.				+	nuclear accent
3.+				+	pitch accent
2.+		+		+	full vowel
1.+	+	+	+	+	syllable

Ronnie loves Marie

Fig. 1 Grid representation of the utterance *Ronnie loves Marie*.

According to this view, there are three degrees of stress prominence above the syllable level, as represented in figure 1 by level 2, 3 and 4. First, this representation allows us to talk about a hierarchy of qualitative distinctions between accented vs. unaccented (though stressed) syllables. For instance, in figure 3 the syllable *loves* is marked by the first level of stress (level 2) because of the nature of its vowel, which is "full" and not reduced, unlike the vowel of the first syllable in *Marie*; however, this syllable is not marked by the second level of stress (level 3), i.e. it is not (pitch-)accented. We must underline that according to this theory of intonational phonology, an accented syllable is the product of the association of a pitch accent with a metrically strong syllable. The third and final level of stress (level 4) defines the nuclear accented syllable, which is the most stressed among all possible accents in the phrase. As we can notice from the grid representation in figure 1, the strongest stress (level 4) is located where prominence, marked by the presence of a cross at each level, reaches the highest point. The pattern that traditional generative theories might have introspected as having "normal stress" is the one depicted above. In these circumstances, the nuclear accent is the last, and considered to be the most prominent pitch accent in a phrase, whereas a pre-nuclear accent is any pitch accent preceding the nuclear one (see the grid mark on *Ronnie* at level 3 in Fig. 1). This representation entails a notion of stress as a complex, structural entity, while previous

notions of stress, typical of older accounts, view stress as a separate, "suprasegmental" feature (see Lehiste, 1970).

It is also clear that prominences can be controlled in order to obtain specific focus effects. A word or a phrase can be singled out by the nuclear accent placement as the most important bit of information that the speaker wants to convey. Accent placement is in fact clearly related to focus of information. The notion of nuclear stress played an important role in the interpretation of the notion of "focus" in studies of generative syntax and semantics of the late 1960s and the 1970s (Chomsky, 1971; Jackendoff, 1972; Selkirk, 1984). In this framework, the nuclear stress was invoked in order to promote the last stress of a "neutral" utterance to the nuclear stress. A neutral utterance is the default case, in which no particular word or phrase in the utterance is singled out as being the most important or the "new information" to convey. The nuclear stress rule was then said to produce "normal stress".

After a long debate about this notion of "normal stress", Ladd (1980) reinstated the validity of the pattern and coined a term for it, that is "broad focus", which is still used today to refer to "focus on whole constituents or whole sentences, not just on individual words" (Ladd 1980, p.3). In this perspective, normal stress describes accent placement when focus is broad. "Narrow" focus is the term that naturally opposes "broad" focus and refers to cases where smaller constituents such as single words are selected as the focused element. One way to think of the notion of focus scope is by considering felicity of a statement as an answer to a particular question. For example, statement B in (1), where the constituent under focus is made of a single word, might be uttered in the context of question A (the word in capital letters indicates the nuclear accented word within the focused constituent marked with brackets).

- (1) A: "Who loves Marie?"  
B: "[RONNIE] loves Marie"
- (2) A: "What does Ronnie feel?"  
B: "Ronnie [loves MARIE]"

However, in (2), where the context is represented again by question A and the answer is B, the constituent under focus is larger, i.e. it is the whole verb phrase.

One of the additional claims of the standard theory of the relationship between focus and accent is that the phonological form of an utterance with late accent placement will be ambiguous between a broad focus interpretation and a late narrow focus interpretation (Chomsky, 1971; Jackendoff, 1972). For instance, a pattern like "Ronnie loves MARIE" is potentially ambiguous between late narrow focus and broad focus readings. In other words, this utterance can be the phonological expression of all of the focus structures in (3).

- (3) a. "[Ronnie loves Marie]<sub>F</sub>"  
b. "Ronnie [loves Marie]<sub>F</sub>"  
c. "Ronnie loves [Marie]<sub>F</sub>"

In (3a) the whole sentence is in focus, in (3b) focus is on the entire verb phrase, and (3c) has narrow focus on the noun.

Accent placement works very similarly in Italian. As in English, the location of a pitch accent associated to a stressed syllable is not fixed in a prosodic phrase. For instance, in a Subject-Verb-Object declarative sentence, accent can be placed on any of the words, according to which of them is in focus. In the sentence *Giovanni ama Maria* "John loves Maria", all of the prominence patterns in (4) are possible.

- (4) a. GIOVANNI *ama* Maria  
 b. Giovanni *AMA* Maria  
 c. Giovanni *ama* *MARIA*

While in (4a) the nuclear accent is early, in (4b) it is medial and in (4c) it is late. According to the standard theory of the relationship between focus and phrasal phonology mentioned above, the accentual pattern of (4c) is ambiguous, since it could signal narrow focus on the object, broad focus on the verb phrase or even broader focus on the sentence as a whole (i.e. completely "new" information). This ambiguity represents a problem for a theory that would directly derive focus from accent placement, and is caused by the lack of a one-to-one correspondence between focus and accent. The question is, then, from the view of perception: "How do listeners determine the breadth of focus, given a certain accent placement?". Under a theory that privileges the role of syntax, such as Selkirk (1984), one derives focus interpretation via argument structure and "percolation" rules that allow the [+F] ("focus") feature specification to percolate up the syntactic tree. Other theories would privilege the semantic contribution and the difference in "informativeness" of broad focused and narrow focused words (see Ladd, 1996, chap. 5, for a review of some other accounts that have been proposed to accommodate them).

The present study tests the predictions of standard Generative Phonology, namely whether late and broad focus are actually perceptually confusable. If the claims of standard phonological theories are true, we predict that utterances with late accent placement, whether having an intended broad focus structure or having late narrow focus, will be equally perceptually ambiguous.

## ACCENT STATUS AND PROMINENCE

The typical prominence structure of a broad focus utterance can be thought of as having the same representation of the utterance in figure 1 (above). When Ladd (1980) formulated the notion of broad focus, though, there was still no clear notion of prenuclear (as opposed to nuclear) accent, since the only accent of the intonation phrase was identified with the nuclear one. However, we know that one of the constraints of the phonology on the prosodic shape of the utterance is that content words preceding the nuclear accented word can often be accented too (see prenuclear accent in figure 1). As mentioned above, current theories of intonational phonology assume that when more than one accent is present in a phrase, it is the last one that will be associated with the designated "most stressed" syllable in the phrase.

Under such a hierarchical view, the prediction is that nuclear accented words will be more prominent than pre-nuclear accented words, which in turn will be more prominent than unaccented words. In this perspective, different types of nuclear accent will be characterized by the same degree of prominence or at least by degrees of prominence that do not reverse the pattern of relative prominence of the prenuclear and nuclear stress within the same phrase. For instance, the nuclear accent typical of English statements is H\*, while a "downstepped" H\* (transcribed as !H\*; see Beckman and Ayers, 1994) tends to be used in a more narrative type of tune and is characterized by a lower fundamental frequency (F0) peak. However, both nuclear accent types will produce equal degrees of perceptual prominence, in that they are attributed an equal status in the stress hierarchy. Previous works on prominence perception had suggested that listeners are able to normalize for the natural F0 declination during the course of an utterance (Pierrehumbert 1979), in that they would perceive a later H\* as being as prominent as a preceding H\*, even though (all other things being equal) the late H\* has necessarily a lower F0 peak than the preceding H\*. Work by Horne (1991) and Ayers (1996), on the other hand, suggests that a nuclear syllable bearing a downstepped !H\* accent may not be as fully prominent as a syllable with a H\* in the same position. The observation is that accent type differences can contribute to the perception of more or less strong prominences.

Accent type differences can be observed also in Italian. The nuclear accent type commonly found in statements with broad focus is characterized by a rather downstepped quality. Figure 3 shows two different productions of the sentence *Mamma andava a ballare da Lalla* "Mom used to go dancing at Lalla's". The upper panel presents an example of broad focus, with a tune transcribed as  $H^* H+L^* L-L\%$ . What is important to notice here is the relatively shallow  $F_0$  variation within the nuclear accented syllable marked with  $H+L^*$ , as opposed to the greater  $F_0$  excursion within the prenuclear accented  $H^*$ . Narrow focus statements differ from broad focus ones in that they present a nuclear pitch accent that is acoustically more salient, and are characterized by a  $H^*+L$  nuclear accent (lower panel).

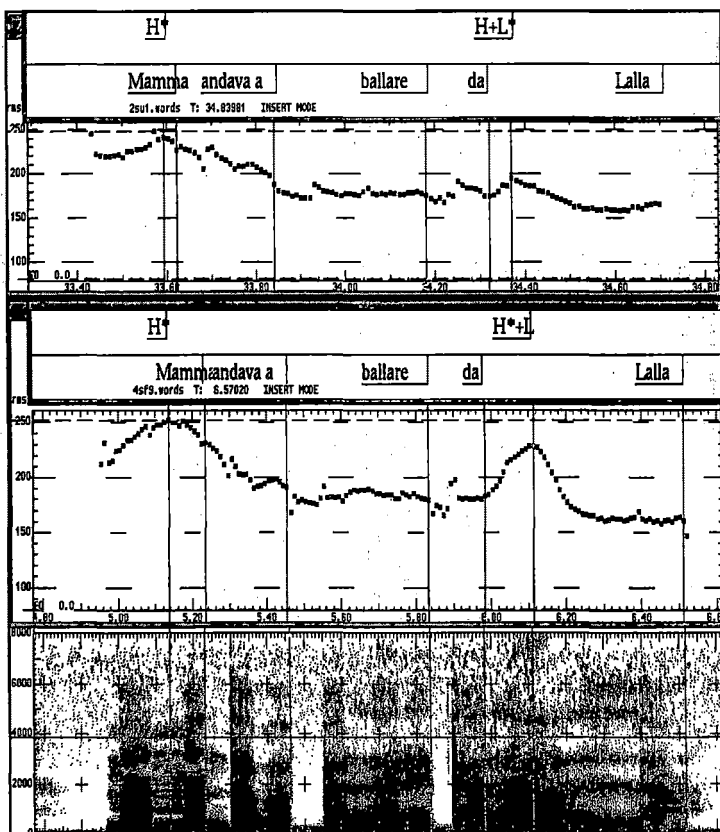


Fig. 2 Broad focus (upper) and late narrow focus (lower) realization of the sentence *Mamma andava a ballare da Lalla* uttered as a statement.

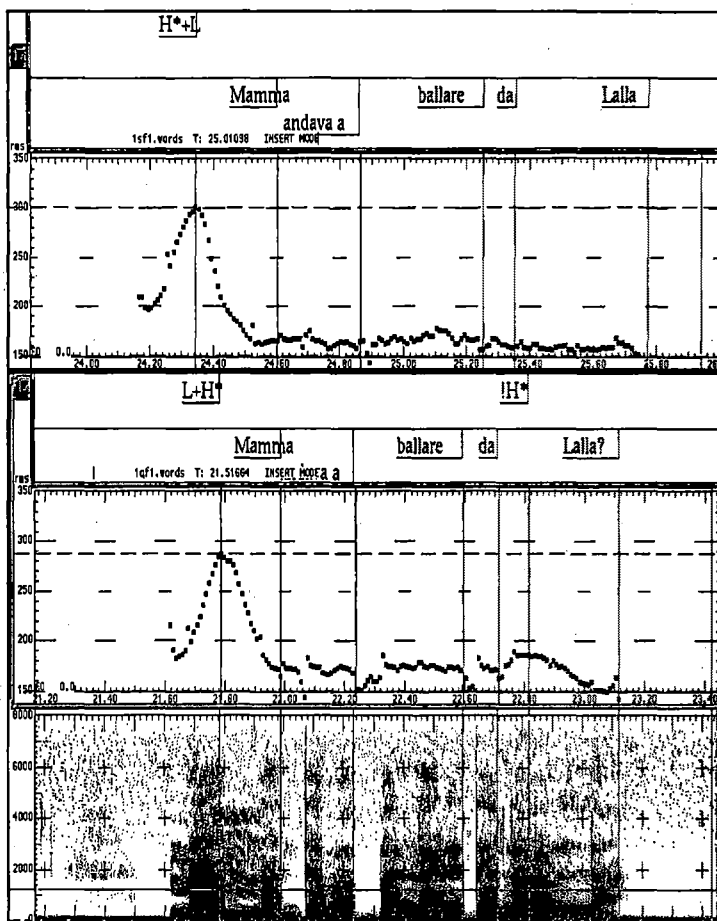


Fig. 3 Pitch contour of MAMMA andava a ballare da Lalla "Mom used to go dancing at Lalla's" uttered with early nuclear accent on *mamma* both as a statement (upper panel) and as a yes/no question (lower panel).

Another important defining feature of the nuclear accent in Pierrehumbert's framework is that it is the pitch accent that immediately precedes the phrase accent in English. A phrase accent is the tonal event that controls the pitch level in the region that goes from the end of nuclear accented word up to the end of the phrase. The role of the phrase tone cannot be underestimated, since it is through this event that we can relate the nuclear accent of English to that of languages like Swedish. While in English the phrase

accent prevents the realization of post-nuclear accents, in Swedish pitch accents occur beyond the phrase accent, but are downstepped (Bruce 1982).

The suggestion that downstepped accents might mark less prominent syllables is especially intriguing when we compare the role of downstep in Swedish and in Neapolitan Italian. While early focus declaratives in both the standard and the Neapolitan variety of Italian present the same characteristics of English early narrow focus utterances (by showing the predicted flat melodic configuration following the nuclear accent), yes/no questions of the Neapolitan variety do not. In this variety, yes/no questions with early focus (see Fig. 3 above) present a sharp pitch obtrusion to a pronounced peak (L+H\*) on the focused constituent, plus a smaller peak on the last stressed syllable of the intonational phrase (IH\*).

To sum up, the standard accounts of nuclear stress in English can be understood as a purely "positional" definition. The nuclear stress is the syllable in D.T.E. position. It is the syllable positioned to hear the last pitch accent in the intonation contour (in English), or the accent positioned just before the phrase accent (in English or Swedish). The common thread underlying all of these definitions is that sentence stress (as represented, e.g., in the grid) is independent of accent type, even if it is not independent of intonation, as assumed in early Generative Phonology. I will refer to this hypothesis as the "positional hypothesis". Moreover, the traditional generative view of the relations between the pragmatics of focus and accent placement claims that an accent placed on the last element of an utterance can ambiguously signal a broad or a late narrow scope of the focus. By the same token, the accent structure of stimuli with either (intended) broad or late narrow focus will be equally perceptually confusable. I will refer to this as the "(late accent) ambiguity hypothesis".

The results of the experiments presented in this paper appear to provide evidence against the ambiguity hypotheses, but we could not find conclusive evidence for the stress-to-prominence relation claimed by the positional hypothesis. Moreover, none of the hypotheses considered could predict the unexpected difference in the patterning of results between questions and statements found here.

## EXPERIMENT I

### METHODS

#### Stimuli

The stimuli consisted of a set of sentences with different word number, focus pattern and modal intonation (question vs. statement). Four groups of sentences, with four sentences in each group, were created. Group I consisted of three-word sentences, uttered as statements; group II consisted of two-word sentences, uttered as statements; groups III and IV were uttered as questions, and included respectively three-word and two-word sentences. All of the sentences had either an SVO (Subject-Verb-Object, e.g. *Maria ama Giovanni* "Maria loves Giovanni") or SV (Subject-Verb, e.g. *Mario esce* "Mario goes out") structure, depending on the number of words in the sentence. The words employed had a variable number of syllables and of lexical stress pattern (initial vs. non-initial).

As shown in Figure 4 (below), each of the sentences for each type was uttered as either a neutral utterance with broad focus (focus B) or as a narrow focused utterance, where scope of focus was limited to a single word - either S (focus S), V (focus V) or O (focus O). The set of sentences was produced by two speakers of Neapolitan Italian (the author and a male speaker), each of whom read half of it. The production of the male speaker was strictly monitored by the author, and the sentences were all auditorily transcribed to check for intended focus pattern.

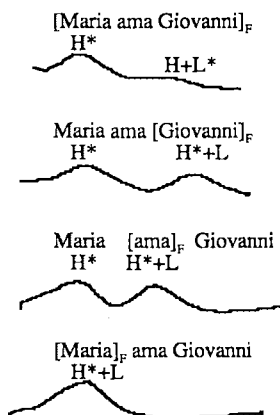


Fig. 4 Stylized statement contours with different focus structure.

The recordings were made in the Department of Linguistics Lab, Ohio State University, where they were digitized at 16 kHz on a SUN Sparc Station using ESPS Waves<sup>†</sup>. The stimuli were coded and pseudo-randomized using a Latin square design. In this arrangement each treatment occurs just once in each row and just once in each column. It was ensured that each of the speakers (the female and the male voice) in the test would not produce two consecutive treatments in each row. The stimuli were then recorded on a tape, which was later played to the subjects.

## Procedure

The test was administered to a homogeneous group of participants, consisting of 20 university students and recent graduates, with ages varying from 23 to 29. From the original 22 participants, the data relative to 2 of them had to be discarded because their response pattern suggested that they were not attending to the task. The participants were all speakers of Neapolitan Italian and hence had the same linguistic background as the speakers that produced the stimuli. Only three of the participants were knowledgeable in Linguistics, but none was aware of the purpose of the experiment.

The method used to gather perceptual data was as follows: each participant was given a reply sheet on which 56 sentences were written, each consisting of either two or three words, separated from each other by means of framed boxes.

Ex.

Mario	esce
-------	------

Maria	ama	Giovanni
-------	-----	----------

The set of sentences was divided into four blocks. The participants listened to each sentence and had to mark only one of the words in the sentence, with a check or a cross on a specific box; specifically, they were asked to mark the word that appeared to be the most "important" in the sentence. Subjects were also told to mark the answer as quickly as possible after listening to each stimulus, even when not entirely sure about the answer. Use of linguistic terms such as "prominence" and "focus" was avoided and the attribute used



throughout the experiment was "important". It was our concern to leave linguistic notions outside of the task, so that even naive listeners could perform it without confusion.

A short training session preceded the set of trials, where the experimenter presented examples of utterances with varying intended focus structure (see Figure 4 above) and had the participant point at one of the words as being the most important. Each sentence was separated from the following by a three-second pause. To facilitate the task, a short beep separated each block of utterances from the other. The beep was inserted to indicate the stage of the experiment reached, while knowing that the fourth beep would signal the end of the experiment itself.

## Design and predictions

Two factors were included in the experimental design. The focus factor, derived from labeling the intended focus pattern following the standard theory, and the modality factor (question vs. statement intonation), will be referred to respectively as FOCUS and MODALITY. FOCUS has 3 levels in the two-word sentences (focus B, focus S and focus V) and 4 levels in the three-word sentences (focus B, focus S, focus V and focus O); MODALITY has always 2 levels (question modality vs. statement modality). Therefore, the design was a 2 (modality) x 3 (focus) factorial in two-word sentences and a 2 (modality) x 4 (focus) factorial in three-word sentences. Both independent variables, FOCUS and MODALITY, were manipulated within subjects.

For two-word stimulus sentences, responses were transformed into the dependent variable PERCENT, which is the percentage of responses which assigned "importance" to the Verb for each stimulus token. We will assume for now that "important" is equal to "prosodically prominent", and on the basis of this assumption we will make the following predictions. According to the positional hypothesis, we expect that in focus B and focus V stimuli, where the word carrying the nuclear accent is the Verb, we will have a mean percent of "assigned importance" of around 100%, whereas the value of percent "importance" assigned to the Subject will be close to 0%. By the same hypothesis, we predict that the value of assigned importance to the Verb will be around 0% in focus S stimuli, while assigned importance to the Subject will be close to 100%. By the ambiguity hypothesis we also predict that, whichever the exact value will be, focus B stimuli will have an exact same value of assigned importance to the Verb as focus V stimuli, since they both have late accent placement.

For three-word stimuli, we have to set three independent variables, since participants had a choice among three and not just two elements for the purpose of assigning importance. The three variables were obtained by calculating percentages of assigned importance to each of the three word positions, i.e. Subject, Verb and Object, yielding respectively the variables PERCENT-S, PERCENT-V and PERCENT-O. Again, according to the positional hypothesis, we expect that, independent of breadth of focus, in all stimuli types (focus B, focus S, focus V and focus O), the word carrying the nuclear accent will have a mean percent of "assigned importance" of around 100%, while the other words will receive values of assigned importance close to 0%. This means, for instance, that the value of PERCENT-S for focus S stimuli will be around ceiling. By the ambiguity hypothesis we again predict that, whichever the exact value, focus B stimuli will have the same value of assigned importance to the Object as focus O stimuli, since they both have late accent placement.

Two-way ANOVAs were performed on the dependent variables. Post-hoc tests (Tukey Kramer) were performed to test the relative significance of the mean differences in all cases of significant interaction. The criterion was set to  $p < .05$  for both planned comparisons and post-hoc tests.

## RESULTS

### Two-word utterances

As expected from the predictions of the positional hypothesis, the highest value of PERCENT-S (95.63%) was found in utterances with narrow focus on the subject. An interesting result is represented by the pattern shown by broad focus utterances, where, contra the predictions, we find a value of 35.63% for PERCENT-S. What is more, utterances with intended focus on the verb received assigned importance on S in 14.38% of the cases, against expected values of around 0%.

In figure 5 the interaction data for both questions and statements are presented. Here, PERCENT-S is plotted on the y axis across the three utterance types (focus S, V and B). Contrary to the expectations, statements with focus on the verb have a mean of 22.5% responses which assigned importance on the subject. Two results were not predicted by either of the theories mentioned above. First, questions received lower overall values for PERCENT-S than statements. Second, there was a conspicuous difference between broad focus questions and statements. While the category of broad focus utterances produced, overall, a high PERCENT-S value, the biggest contribution is due to broad focus statements, which averaged 55% PERCENT-S (as compared to 36.5% PERCENT-S for questions).

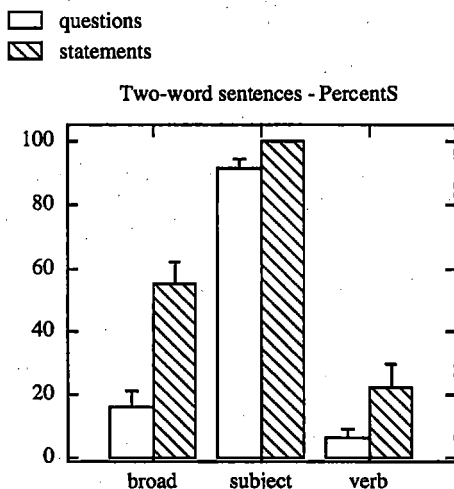


Fig. 5 PERCENT-S values by FOCUS (Broad, Subject, Verb) and MODALITY (question vs. statement) for two-word utterances.

Both FOCUS [ $F(2, 18) = 136.0$ ;  $p < .05$ ] and MODALITY [ $F(1, 18) = 25.924$ ;  $p < .05$ ] were found to significantly influence the pattern of responses. A significant interaction between the independent variables was also found [ $F(2, 18) = 4.664$ ,  $p < .05$ ]. A post-hoc comparison revealed a significant difference between all the levels of the first and the second independent variable. Therefore, the observed trend of utterance modality (question vs. statement) in determining the percentage of assigned importance was proved to be significant.

### Three-word utterances

#### Narrow focus utterances

Figure 6 shows the pattern of PERCENT-S responses for three-word sentences. In utterances with narrow focus, most subjects judged the focused word as most important, without regard to modality, with 94% of responses judging the Subject as "most important" in focus S utterances, 87% judging the Verb as most important in focus V utterances and 77% judging the Object as most important in focus O utterances. Again, with the assumption that prominent equals "important", we can make various observations. In a way that was not predicted by the positional hypothesis, utterances with narrow focus on the object received a fair amount of assigned importance on the subject, i.e. 17%, and a minor percent of assigned importance on the verb, i.e. 6%. In utterances with narrow focus on the verb, just 6% of the subjects judged the intended focus to be on the subject and 7% to be on the object.

Figure 6 reveals also that modality, i.e. question vs. statement intonation, affects the pattern of utterances with narrow focus on the verb. In particular, questions produce a percentage of 79% for this intended focus structure, whereas statements yield a value of 95%. This trend can be found, to a lesser extent, in utterances with focus on the subject; within this category, questions produced 89% of PERCENT-S responses, against 99 % for statements.

As for the statistics on the values of PERCENT-S, both FOCUS [ $F(3, 24) = 71.220$ ;  $p < .05$ ] and MODALITY [ $F(1, 24) = 4.854$ ;  $p < .05$ ] resulted significant. Similarly, an interaction of the independent variables was found [ $F(3, 24) = 3.523$ ;  $p < .05$ ]. Post-hoc tests revealed significant differences between utterances with focus S and every other focus level. The ANOVA performed on data for PERCENT-V revealed a significant effect of FOCUS [ $F(3, 24) = 136.00$ ;  $p < .05$ ], once again, but neither a significant difference of MODALITY [ $F(1, 24) = 1.05$ ;  $p > .05$ ] nor a significant interaction [ $F(3, 24) = 2.235$ ;  $p > .05$ ]. A post-hoc comparison uncovered significant differences between utterances with focus on the verb and every other focus level. The data of assigned importance to the object will be discussed below.

#### Broad focus utterances

For broad focused utterances, the predictions of the positional hypothesis are not met. In fact, we find that only 56% of the responses assigned importance to the object within this focus category, as opposed to the very high value of this percentage for utterances with narrow focus on the object. Once we assume that importance equals prominence, the prediction of current phonological theory is that these two percentages should be roughly equal, since there should be no phonological prominence difference between utterances with late accent placement, whether they be narrowly or broadly focused. It was also found that the subject and the verb received a fair amount of assigned importance when focus was broad, in the measure of 26% and 18% respectively.

Even more interestingly, when looking at figure 6, we notice again that broad focus utterances pattern differently according to modality. First, the interaction pattern is reversed relative to what we saw previously. Overall, statements present smaller percentage values of assigned prominence to the object than questions. Remarkably, broad focus statements present just 35% of assigned importance to the object. As observed above, statements with narrow focus on the object presented instead a much higher value, i.e. 77.5%. Questions, on the contrary, produced an equal percentage of responses for both of these focus levels, (76%), giving support to the ambiguity hypothesis. Moreover, broad focus statements present a very high and unexpected percent of assigned importance to the subject (43.74%), whereas broad focus questions produced a much smaller value (8.75%). The

data relative to broad focus statements also seem to possess the greatest variability when compared to data for other focus structure types.

The analysis of variance on the percent of assigned importance to the object uncovered a significant effect of FOCUS [ $F(3,24) = 47.977$ ;  $p < .05$ ] and an effect of MODALITY [ $F(3,24) = 6.691$ ;  $p < .05$ ]. The interaction was also significant [ $F(3,24) = 3.207$ ;  $p < .05$ ]. A post-hoc test on the focus effect revealed a significant difference between focus B and focus S, focus B and focus V, focus O and focus S and focus O and focus V utterances. A post-hoc analysis performed on the modality effect resulted in a significant difference between questions and statements.

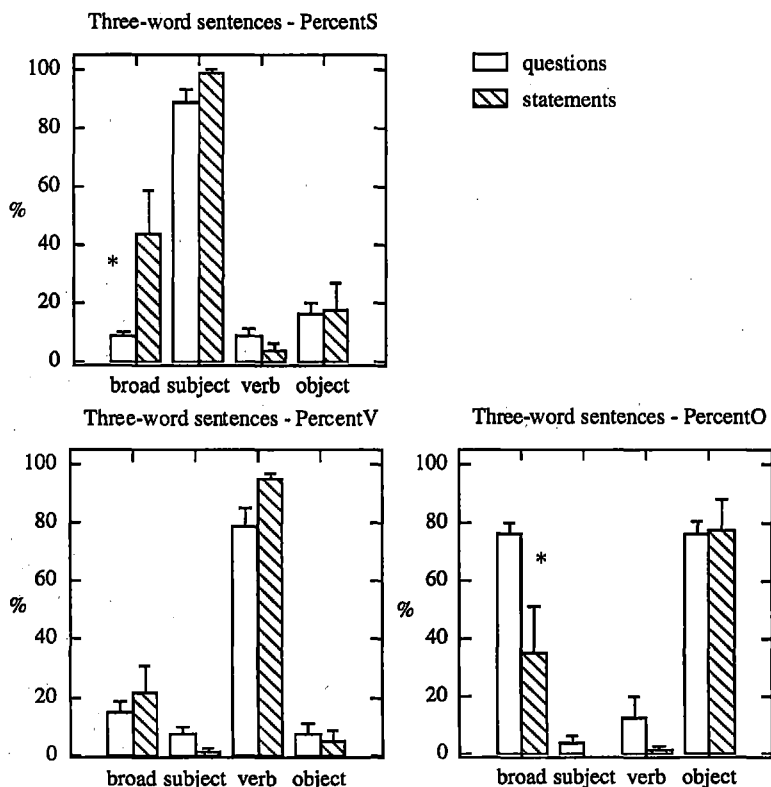


Fig. 6 PERCENT-S (upper), PERCENT-V (medial, left) and PERCENT-O (medial, right) values by FOCUS (Broad, Subject, Verb) and MODALITY (question vs. statement) for three-word utterances.

Therefore, contrary to the predictions of current phonological theory, a significant difference was found in the patterning of the data between utterances with broad focus and

narrow focus on the object. As a relevant and novel result, while focus B and focus O questions produced identical values of PERCENT-O, focus B and focus O statements did not.

## DISCUSSION

Experiment I attempted to assess the percept of prosodic prominence through the notion of "importance". However, the results are difficult to interpret in this light, and seem to suggest that listeners were responding in terms of focus structure, and not simply in terms of accent structure. Moreover, while a traditional view of the relationship between accent placement and focus structure (the ambiguity hypothesis) predicted that late accent utterances in general would show identical responses for both broad and late narrow focus, our results show that this is true only for questions.

The most interesting result of Experiment I, in fact, comes from the analysis of broad focused utterances as a function of modal intonation. Specifically, while statements with intended broad focus (focus B) obtained values of assigned importance that are significantly different from those obtained for statements with intended late narrow focus (focus V for two-word stimuli and focus O for three-word stimuli), the converse was true for questions.

To sum up, the results of Experiment I suggest that the listeners relied on the focus structure of the utterance more than on its prominence structure (given by the relation between nuclear and prenuclear accent in a prosodic hierarchy of accent status) in order to perform the "assigned importance" task. This unexpected outcome urged us to perform a second experiment where the notion of focus will be investigated more directly, in order to see to what extent listeners were responding in terms of focus in Experiment I and not in terms of accent relationships. However, Experiment II will still leave unanswered the question relative to prominence perception, which will probably need the use of a different experimental methodology. An alternative hypothesis for the unexpected results of the broad focus statements is that listeners were simply confused when it came to assigning "importance" to a word in an utterance with no prior context. Experiment II will also try to address this concern.

## EXPERIMENT II

### METHODS

#### Stimuli

The same stimuli as for Experiment I were employed (see the Stimuli section for Experiment I).

#### Procedure

The test was administered to a homogeneous group of participants, consisting of 23 university students and recent graduates, with ages varying between 22 and 29, and who were all speakers of the Neapolitan variety of Italian. Participants had to match the statement they heard to one of several questions setting up a specific context:

Ex. stimulus: *Maria ama GIOVANNI* "Mary loves JOHN"  
matching question: *Chi ama Maria?* "Who does Maria love?" (focus O)  
other choices: *Chi ama Giovanni?* "Who loves John?" (focus S); *Che pensa Giovanni di Maria?* "What does Maria think of John?" (focus V) or *Dimmi qualcosa di quella coppia* "Tell me something about that couple" (focus B).

When the stimulus heard was a question, the set of context-setting choices was very similar to the one used for statements, but in this case the participants were instructed to choose the question whose meaning most resembled that of the stimulus question, or that was felicitous as an utterance following what they just heard. For instance, in a dialogue, the question *Chi ama Maria* "Who loves Maria?" is felicitous in the context of an immediately following yes/no question *GIOVANNI ama Maria?* "John loves Maria?".

Participants first listened to the test sentence and then chose the context question.

## Design and predictions

The experimental design was the same as for Experiment I. The only difference was in the number of dependent variables on which the statistics was conducted. In Experiment II, the participants had to decide between three possible matching questions for two-word utterances and four for three-word utterances. The coding devised for each question typology was the following: PERCENT-B stands for percent of responses choosing the question congruent with a broad focus utterance, while PERCENT-S, PERCENT-V and PERCENT-O refer respectively to percent of responses choosing questions congruent with a narrow focus utterance (either focus S, focus V or focus O).

As in Experiment I, given the hypothesis that participants were responding in terms of focus structure, we expect very high values of PERCENT-S, PERCENT-V and PERCENT-O responses for utterances with narrow focus on the respective element. Especially high values are expected for questions and two-word utterances. Since the participants had to match a specific stimulus to a matching context-question, we also expect them to be more consistent in the pattern of responses than in Experiment I. This factor could especially be relevant for broad focus utterances, which produced highly variable responses in Experiment I. Our hypothesis can be formulated as follows:

Hypothesis 1: following the ambiguity hypothesis of traditional focus-to-accent accounts, we expect ambiguity of focus extraction for focus B and late narrow focus utterances (focus V for two-word stimuli and focus O for three-word stimuli). We must recall that this ambiguity was suggested by the results of Experiment I to be very high in question stimuli. Therefore, we expect high values of PERCENT-V for two-word broad focus utterances, where late accent placement will simply be confused with late narrow focus, and, by the same token, high PERCENT-O values for three-word broad focus utterances. This hypothesis predicts also that high values of PERCENT-S and PERCENT-V will be found for early narrow focus utterances, while late narrow focus utterances will receive equal percentages of PERCENT-O and PERCENT-B. This hypothesis is based on the claim that accent placement is what renders broad focus ambiguous and virtually indistinguishable from late narrow focus.

On the other hand, given the results of Experiment I, if hypothesis 1 is not completely supported it will be that broad focus statements will not necessarily be identified with late narrow focus statements, while broad focus questions will. Broad focus statements will instead present high values of PERCENT-B responses. Given the value of assigned importance to the Subject for broad focus statements in Experiment I, it could also be the case that broad focus statements will receive focus S responses in a few cases. If this scenario is confirmed by the results of Experiment II, we will have to find a suitable explanation.

## RESULTS

### Two-word utterances

#### Narrow focus utterances

The results clearly replicated those of Experiment I. The highest value of PERCENT-S (93%) was found in utterances with narrow focus on the subject, while the highest value of PERCENT-V (81%) was found in utterances with narrow focus on the Verb.

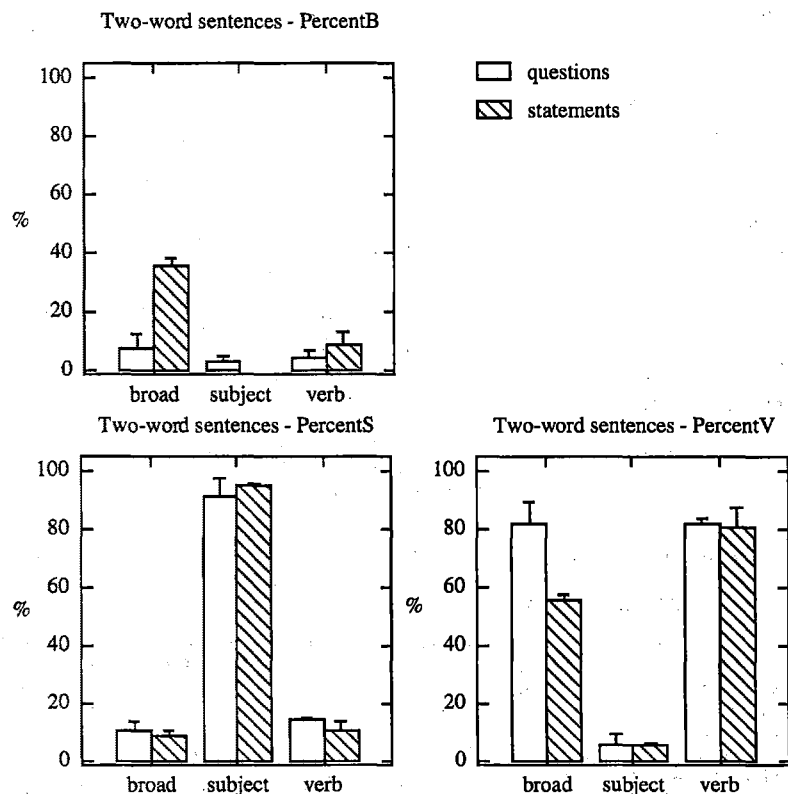


Fig. 7 PERCENT-B (upper), PERCENT-S (medial, left) and PERCENT-V (medial, right) values by FOCUS (Broad, Subject, Verb) and MODALITY (question vs. statement) for two-word utterances.

The ANOVA performed on PERCENT-S yielded the following results: FOCUS affected significantly the pattern of responses [ $F(2,18) = 453$ ;  $p < .05$ ], while MODALITY [ $F(1,18) = .80$ ;  $p > .05$ ] and the interaction between the variables [ $F(2,18) = .620$ ;  $p > .05$ ] did not. As for PERCENT-V, FOCUS was again highly significant [ $F(2,18) = 147$ ;  $p < .05$ ], but there was also a significant effect of MODALITY [ $F(1,18) = 5.531$ ;  $p < .05$ ] and a significant interaction [ $F(2,18) = 4.894$ ;  $p < .05$ ].

In figure 7 the interaction data for both questions and statements are presented. Here, PERCENT-B, PERCENT-S and PERCENT-V are plotted on the y axis across the three utterance types (focus S, V and B). Contrary to the expectations, statements with focus on the verb have a mean of 10% responses which assigned importance on the subject. However, this value is lower than the one reported for the same focus type in Experiment I. Also, differently from Experiment I, questions did not receive a lower overall value of PERCENT-S responses.

### **Broad focus utterances**

Figure 7 shows also results for broad focus utterances. The results of Experiment I are replicated also here. The value of PERCENT-V is once again lower than the value reported for utterances with narrow focus on the Verb. As in Experiment I, broad focus statements receive a lower score of PERCENT-V than late narrow focus ones, even though the value is well above chance here (55%), which appears to support hypothesis 1. Also, the value of PERCENT-S is much smaller than the one reported for Experiment I (around 10%). Broad focus and focus V questions receive same values of PERCENT-V, strongly supporting hypothesis 1.

As to the values of PERCENT-B, they are highest for broad focus utterances, as expected, even though this value is relatively low (22%, below chance). Figure 7 reveals also a highly significant interaction [ $F(2,18) = 13.471$ ;  $p < .05$ ] between the two modalities. The value of PERCENT-B is due mainly to the effect of broad focus statements, which alone present a PERCENT-B value that is slightly above chance (36%). FOCUS was also significant in this case [ $F(2,18) = 21.902$ ;  $p < .05$ ], as well as MODALITY [ $F(2,18) = 14.294$ ;  $p < .05$ ]. A post-hoc comparison revealed a significant difference between questions and statements, as well as between broad focus and the other two focus types. The results confirm the tendency for broad focus questions to be identified with late narrow focus which was already noticed in Experiment I.

### **Three-word utterances**

#### **Narrow focus utterances**

Figure 8 shows the pattern of PERCENT-S, PERCENT-V and PERCENT-O for three-word sentences with narrow focus, which replicated the results of Experiment I. In utterances with narrow focus, most subjects judged the focused word as most important, without regard to modality, with 92% of responses judging the Subject as "most important" in focus S utterances, 76% judging the Verb as most important in focus V utterances and 73% judging the Object as most important in focus O utterances. Differently from Experiment I, utterances with narrow focus on O were not incorrectly identified with focus S utterances in many cases (only 7%), while they were identified with focus V utterances with a slightly higher percentage (17%). In utterances with narrow focus on V, just 10% of the subjects judged the intended focus to be on S and 12% to be on O.

The percentage of subjects judging the S as being most important was significantly different only by FOCUS [ $F(3,24) = 154$ ;  $p < .05$ ], while no effect of MODALITY [ $F(1,24) = .128$ ;  $p > .05$ ] nor any interaction [ $F(3,24) = .301$ ;  $p > .05$ ] were found. Post-hoc tests revealed significant differences between utterances with narrow focus on the subject and every other focus level. The ANOVA performed on PERCENT-V data revealed again only a



main effect of FOCUS [ $F(3,24) = 98.7$ ;  $p < .05$ ], and neither a significant difference due to MODALITY [ $F(1,24) = .803$ ;  $p > .05$ ] nor a significant interaction [ $F(3,24) = .93$ ;  $p > .05$ ]. A post-hoc comparison uncovered significant differences between utterances with focus on V and every other focus level. The data for focus O and broad focus utterances will be discussed in the section below.

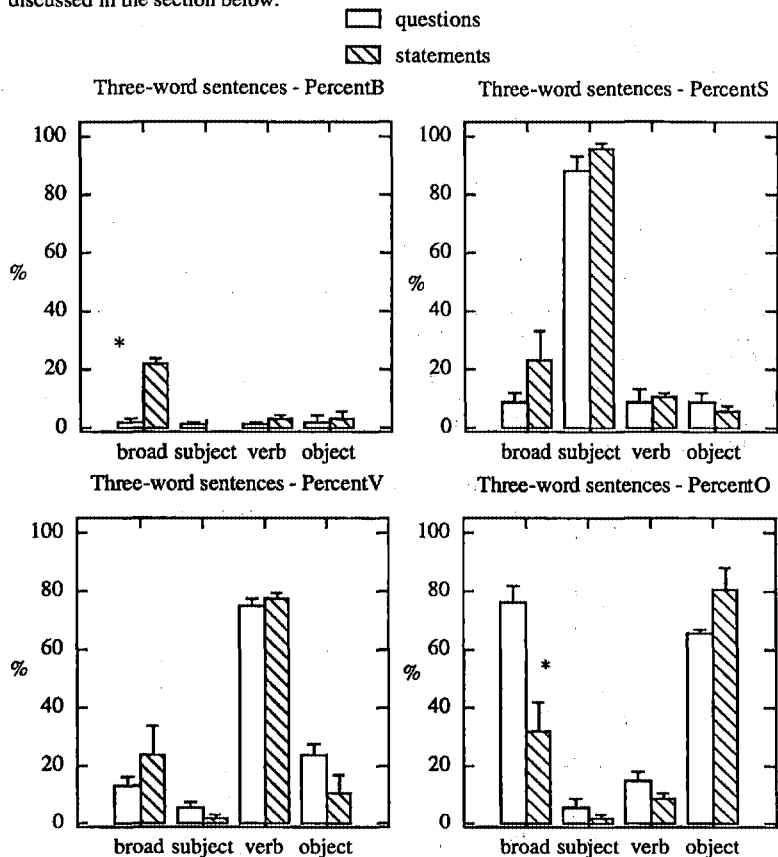


Fig. 8 PERCENT-B (upper left), PERCENT-S (upper right), PERCENT-V (medial, left) and PERCENT-O (medial, right) values by FOCUS (Broad, Subject, Verb) and MODALITY (question vs. statement) for three-word utterances.

### Broad focus utterances

For three-word broad focus utterances, the predictions of hypothesis 1 are not completely met. Again, we find that only few (12%) of the responses matched broad focus

utterances with questions that elicit narrow focus on O, as opposed to the much higher percentage for utterances with intended narrow focus on O. The prediction of hypothesis 1 is, instead, that these two percentages should be roughly equal, since there is no difference in accent placement between utterances with intended late narrow focus and broad focus.

Upon closer inspection, we notice once again that the effect is due in large part to broad focus statements alone (22%). Indeed, the ANOVA performed on PERCENT-O revealed a significant interaction of the independent variables [ $F(3,24) = 11.622$ ;  $p < .05$ ] and a significant effect of MODALITY [ $F(1,24) = 7.07$ ;  $p < .05$ ], in addition to the expected effect of FOCUS [ $F(3,24) = 80.94$ ;  $p < .05$ ]. A post-hoc test on the focus effect revealed, as in Experiment I, a significant difference between focus B and focus S, focus B and focus V, focus O and focus S and focus O and focus V utterances.

The ANOVA on PERCENT-B revealed a highly significant effect of both MODALITY [ $F(1,24) = 27.27$ ;  $p < .05$ ] and FOCUS [ $F(3,24) = 24.54$ ;  $p < .05$ ], in addition to a highly significant interaction [ $F(3,24) = 20.91$ ;  $p < .05$ ]. Broad focus questions were almost never identified as such, while being associated primarily to questions which are congruent with a focus O question. It is interesting that broad focus questions received higher values of PERCENT-O than questions with narrow focus on O. Figure 9 illustrates the highly ambiguous results for broad focus statements, with values of PERCENT-B, PERCENT-S, PERCENT-V and PERCENT-O juxtaposed. Almost equal values of each of the various response percentages are found for this utterance typology, with PERCENT-O being the highest and PERCENT-B the lowest.

Contrary to previous results, statements with narrow focus on the object presented higher percentage values of focus identified on the object than questions. The data for broad focus statements presented again the greatest variability, at least within the three-word utterance group.

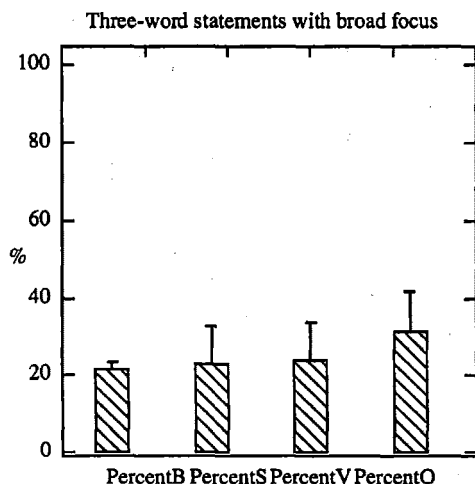


Fig. 9 PERCENT-B, PERCENT-S, PERCENT-V and PERCENT-O for broad focus statements.

## GENERAL DISCUSSION

The results of Experiment II remarkably replicate those of Experiment I. For broad focus results in Experiment II, we need to distinguish two-word utterances from three-word utterances. Responses for two-word utterances can be broken down into two groups, in that almost equal PERCENT-V values were found for both broad focus and focus V utterances as a whole, with little difference between questions and statements. These data appeared to support Hypothesis 1 by revealing ambiguity in the identification of breath of focus when accent placement is late. Under this view, broad focus utterances and focus V utterances are both reported by listeners as having intended narrow focus on the last element. The same cannot be said for PERCENT-B results, however. Broad focus utterances produce a much higher value of PERCENT-B than focus V utterances, where the major contribution is due to statements. Also, the value of PERCENT-S for broad focus utterances was almost irrelevant in focus B utterances as a whole, as opposed to the high value reported for the same focus type in Experiment I.

Three-word broad focus utterances presented an overall pattern of responses which is much more comparable to the one found in Experiment I. Three-word statements with focus B produced lower PERCENT-O (which is the latest focus placement in three-word utterances) responses than utterances with focus O. This was not the case for questions, where focus O and focus B utterances patterned identically. Therefore, as a replica of Experiment I, while focus B and focus O questions produced an almost identical pattern of PERCENT-O responses, focus B and focus O statements did not. PERCENT-B values were once again restricted to broad focus statements. Differently from the results of two-word stimuli, broad focus statements produced a considerable amount of responses where a focus S structure is identified.

It is very likely that the question-matching task taps into the notion of focus in a more reliable way than the "assigned importance" task, which is supported by the less variable pattern of overall responses in Experiment II. However, despite the new task, three-word utterances present a picture that highly resembles that of Experiment I. Three-word broad focus statements are almost equally identified as having narrow focus on the subject noun, on the verb, on the object (which obtains slightly higher values relative to the other responses) or as being broadly focused. What is more, the "wide scope" ambiguity appears to be restricted to statements alone, while questions present ambiguity restricted to broad focus and late narrow focus parse.

Overall, the results appear to match our expectations from the outcome of Experiment I, in that we replicated the confusion for broad and late narrow focus questions, as well as the higher uncertainty in the focus structure of broad focus statements. This can be explained with the observed differences in accent type between questions and statements, on one side, as well as between narrow focus and broad focus on the other. As we mentioned above, the "special" status of downstepped accents has been noted for English in the psycholinguistic study of Ayers (1996). Additional evidence for the special status of downstepped pitch accents comes from ambiguity in prosodic transcription: transcribers either disagree in the placement of !H\* or fail to recognize its presence in the utterance altogether (Pitrelli et al., 1994; Ross, 1995 cited in Beckman, 1996). Broad focus statements in Neapolitan Italian possess a nuclear accent that has a downstepped quality and is very different from the nuclear accent of narrow focus statements (see Figure 2 above). This could explain the uncertainty on the part of the speakers in focus extraction tasks. The peculiarity of broad focus has also been recently noticed in non-Western languages (see Jin, 1996).

Accent type considerations might predict that the acoustically salient L+H\* of Neapolitan Italian questions will facilitate the task of focus extraction for this modality. As we saw from the results of both Experiment I and II, early focus questions, where the nuclear accent is followed by a non-prominent post-nuclear accent, are always correctly identified as having an intended early focus structure. This suggests that a simple positional

definition of nuclear stress prominence is not enough and that accent type considerations will have to be formalized to integrate the theory. Such an integrated theory of stress prominence will correctly predict lower ambiguity for accent structures expressing narrow focus scope, because these will be characterized by a phonetically and phonologically more prominent accent type. Similarly, we showed that focus interpretation does not seem to be guided simply by accent placement, but also by choice of accent type. A simplistic theory of the relations between the semantics of focus and the phonology of accent structure that predicts the same parsing difficulty for intended broad focus and late narrow focus will be bound to fail, while an alternative view should predict easier parse for narrow focus structure, where it is relatively simple to retrieve the scope of focus since it is more restricted.

Certainly, the problem of the relation between accent status and prominence on one side and accent structure and focus extraction on the other is particularly intricate for Italian. While it goes with no doubt that Italian possesses some kind of sentence stress, we are still not sure about the constraints operating on it. Among the other things, we still do not know if the sentence stress, or nuclear accent, is, as in English, the pitch-accent that immediately precedes the phrase accent, or if a different characterization is needed. Future research will try to address this issues and will also need to find tasks that will tap into the notion of perceptual prominence in a more reliable way.

## CONCLUSION

The present study shows that the predictions of standard generative theories by which broad focus and late narrow focus utterances are confused with each other is true for Neapolitan Italian in the case of questions. However, broad focus statements present a more complex picture, which might be explained by invoking semantic and accent type considerations. On one hand, when focus is broad, i.e. when the focused constituent is larger, the task of parsing focus is more difficult. On the other hand, inherent acoustic prominence due to accent type differences might also play a role, in that the nuclear accent of narrow focus utterances is markedly different from the one that characterizes broad focus utterances.

We also showed that, when appropriate context is presented to the listeners, intended focus can be more reliably identified. The ambiguity of late nuclear accent questions and statements, however, still remains, regardless of the nature of the task. If accent type is, as it seems, an important factor in prominence perception, it will be highly interesting to study what acoustic characteristics are influential in determining more or less prominent accent structures and the interplay of acoustic cues and modality.

## REFERENCES

- Ayers, G. (1996). *Nuclear Accent Types and Prominence: Some Psycholinguistic Experiments*. OSU Dissertations in Linguistics.
- Beckman, M.E. (1986). *Stress and Non-stress Accent*. Dordrecht, Foris Publications.
- Beckman, M.E. (1996). The Parsing of Prosody. *Language and Cognitive Processes*, 11, pp. 17-67.
- Beckman, M.E. & Ayers, G. (1994). *Guidelines for ToBI Labeling*, vers. 2.0, Ms., Ohio State University.
- Bruce, G. (1982). Developing the Swedish intonation model. *Working Papers* 22, Department of Linguistics and Phonetics, Lund University, pp. 51-116.
- Chomsky, N. and Halle, M. (1968). *The sound pattern of English*. New York, Harper & Row.
- Chomsky, N. (1971). Deep Structure, Surface Structure, and Semantic Interpretation. D. Steinberg and L. Jakobovits (eds.), *Semantics: an Interdisciplinary Reader in Philosophy, Linguistics and Psychology*. Cambridge, CUP, pp. 183-216.

- Crystal, D. (1969). *Prosodic systems and intonation in English*. Cambridge, CUP.
- Fry, D.B. (1958). Experiments in the perception of stress. *Language and Speech*, 1, pp. 126-152
- Halliday, M.A.K (1967). *Intonation and grammar in British English*. The Hague, Mouton.
- Horne, M. (1991). Phonetic correlates of the new/given parameter. *Proceedings of ICPHS '91*, Aix-en Provence, 5, pp. 230-233.
- Jackendoff, R. (1972). *Semantic interpretation in generative grammar*. Cambridge, MA, MIT Press.
- Jin, S. (1996). *An Acoustic Study of Sentence Stress in Mandarin Chinese*. Ph.D. dissertation, OSU.
- Ladd, R.D (1980). *The structure of intonational meaning*. Bloomington, Indiana University Press.
- Ladd, R.D (1996). *Intonational Phonology*. Cambridge, CUP.
- Lehiste, I. (1970). *Suprasegmentals*. MIT, Cambridge, MA.
- Liberman, M. and Prince, A. (1977). On stress and linguistic rhythm. *Linguistic Inquiry*, 8, pp. 249-336.
- Pierrehumbert, J. (1979). The perception of fundamental frequency declination. *JASA*, 66, pp. 363-9.
- Pierrehumbert, J. (1980). *The phonology and phonetics of English intonation*. Unpublished Ph.D. thesis, MIT.
- Pierrehumbert, J. (1993). Alignment and Prosodic Heads. A. Kathol and M. Bernstein (eds.), *ESCOL '93*, pp. 268-286.
- Selkirk, E. (1984). *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge, MA, MIT Press.

